

Technical Assignment 4

Zhuoying Li

1004021202

Overview & Research Question

Our research question for this project is: Is there a significant interaction between various groups of individuals (Nondemented, Demented, Converted), time, and the MMSE score?

Exploratory Data Analysis

Before we dive deep into our research questions, some basic exploratory data analysis is conducted to have an initial understanding of the dataset at hand.

First, we computed some generic summary statistics for our data.

	Age	EDUC	SES	MMSE	CDR	eTIV	nWBV	ASF
Count	76.41	14.56	2.49	27.26	0.30	1478.85	0.73	1.20
Min	60.00	6.00	1.00	15.00	0.00	1106.00	0.64	0.88
Max	98.00	23.00	5.00	30.00	2.00	2004.00	0.83	1.59

Figure 1: Summary Statistics

Now checking the null values for each column, we see that SES contains 15 null rows, and MMSE contains 1 null row. We fix this by filling the null rows with the means.

Drawing the bar graph of number of patients in each group, we get the following result.

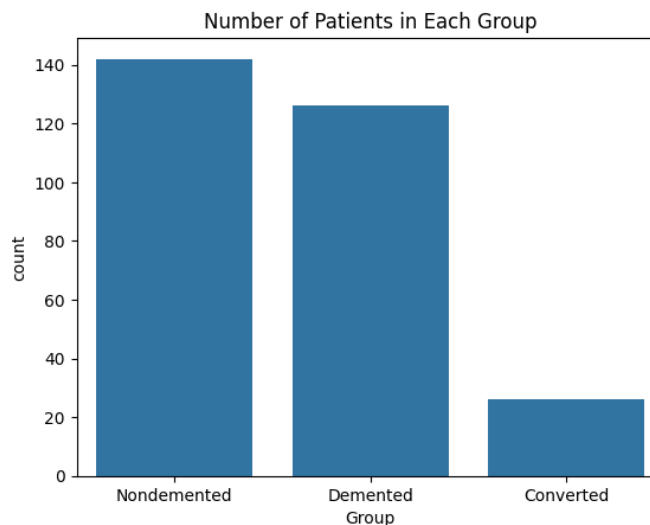


Figure 2: Number of Patients in Each Group

We see a significant difference between number of patients in each group, especially for converted, we have very little sample, which can lead to issues with power down the line.

Drawing a histogram of MMSE scores, we get the following graph.

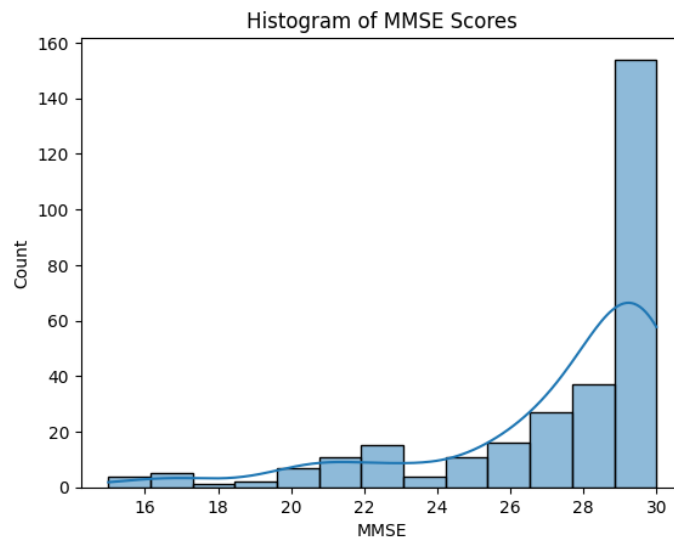


Figure 3: Histogram of MMSE Scores

Looking at the histogram, we can see the MMSE scores is highly skewed, doesn't follow a normal distribution at all.

Because we are interested in seeing the effects of groups on the MMSE score, we draw a pair-wise boxplot, grouped by different groups, each separated by visits.

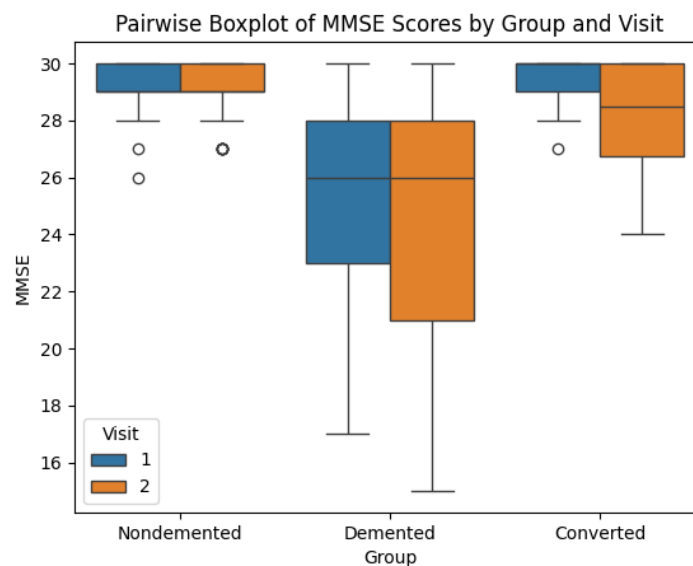


Figure 4: Pairwise Boxplot of MMSE Scores by Group and Visit

From the graph, we can see that overall, nondemented group's MMSE score doesn't seem to have any different between visits, while converted group's MMSE score have a significant difference, where the second visit tends to have a lower MMSE score.

Drawing the interaction plot also provides us a similar conclusion.

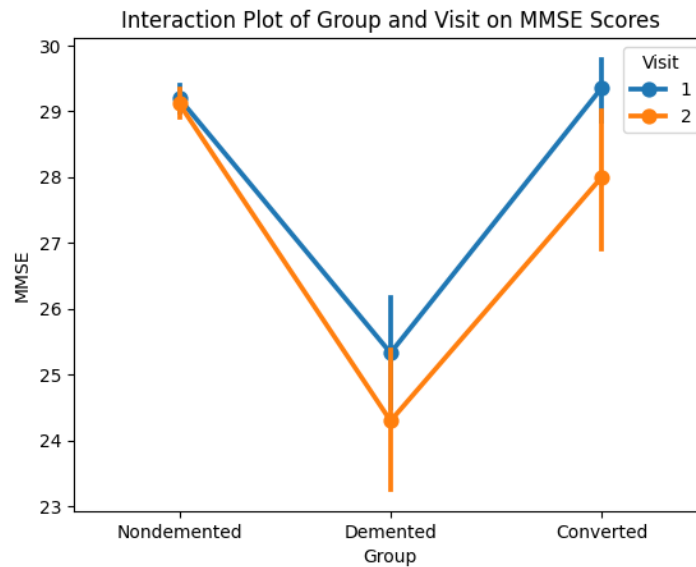


Figure 5: Interaction Plot of Group and Visit on MMSE Scores

Mixed-Effects ANOVA

Given our research question, we form the following hypothesis

- H_0 : There is no significant difference between groups across visits for the MMSE score.
- H_1 : There is a significant difference between groups across visits for the MMSE score.

We constructed the mixed-effect ANOVA model using the following parameters

- Dependent Variable: MMSE
- Within Subjects Factor: Visit
- Between Subjects Factor: Group

After running the ANOVA model, we obtain the following results

Source	P-Value	DF
Group	0.000 (< 0.001)	2
Visit	0.004	1
Interaction	0.043	2

Figure 5: ANOVA Table

Reading the table, we can see that our between subjects factor (Group) has a significant impact on MMSE score (p-value < 0.001). While the within subjects factor (Visit) doesn't have any significant impact (p-value 0.004), same goes for interaction (p-value 0.043).

Assumption Checking

Now we have our mixed-effect ANOVA results, we must check for assumptions to ensure that our results are reliable.

The first thing to check is the normality of residuals, the result is shown in the following table.

Visit	W	Pval	Normal
1	0.787462	1.761319e-13	No
2	0.761123	4.942098e-14	No

Figure 6: Normality of Residuals

The table shows that both visits have $p\text{-val} < 0.001$, indicating the violation of this assumption.

Second assumption is the homogeneity of variances, for this, we do a Levene's test and the result is shown below.

	W	Pval	Equal Var
Levene	2.200486	0.139046	Yes

Figure 7: Result of Levene's Test

Since the $p\text{-value}$ is > 0.001 , our data satisfies the homogeneity of variances assumption.

Power Analysis

After conducting a power analysis, we found 45 samples to be the most appropriate sample size to achieve power of 0.91, with alpha 0.05 and effect size of 0.7. This can be observed in the graph below.

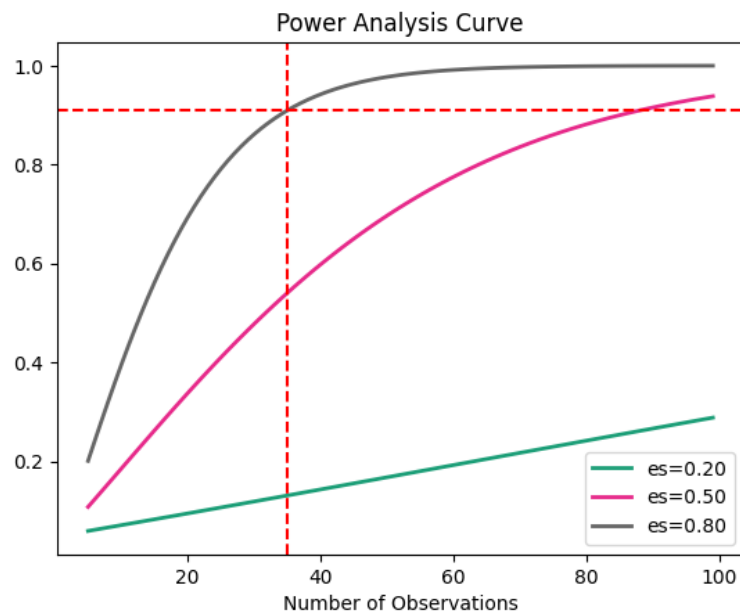


Figure 8: Power Analysis Curve

The sample size we have used is a lot larger than 45, indicating a good statistical power of our results.